

# 海量数据搬运工Datastream

顾费勇

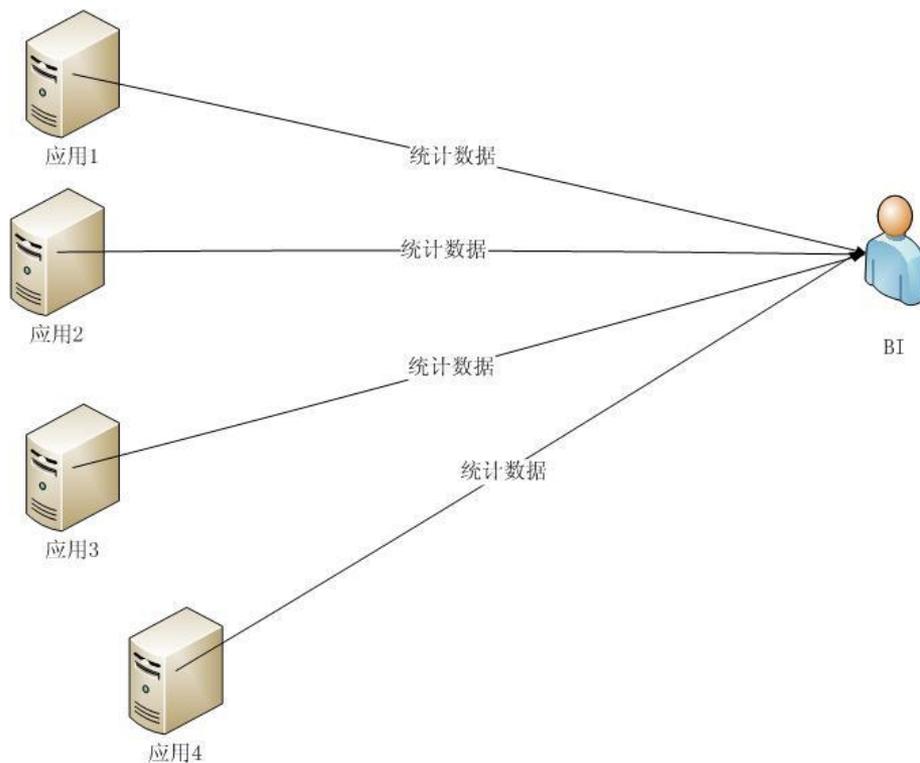
2012-11-28

# 大纲

- Datastream产生的原因
- Datastream的结构和特征
- Datastream关键技术点分享
- Datastream应用场景
- Datastream未来展望

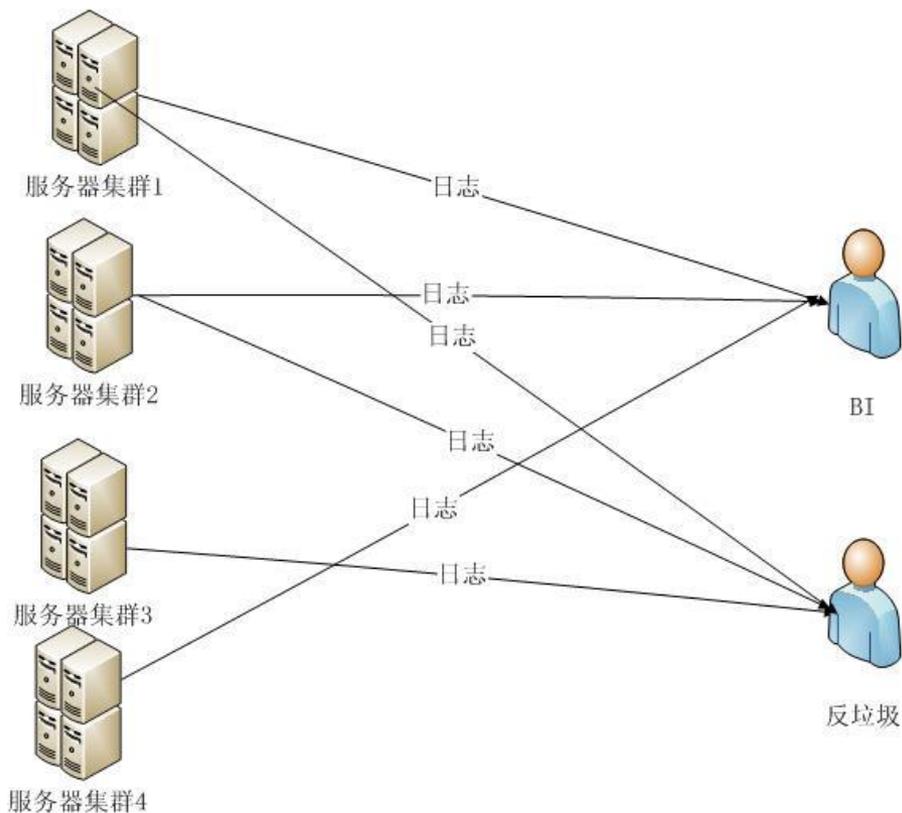
# 场景1

- BI, xx应用想做下日志的统计分析, 收集一下吧



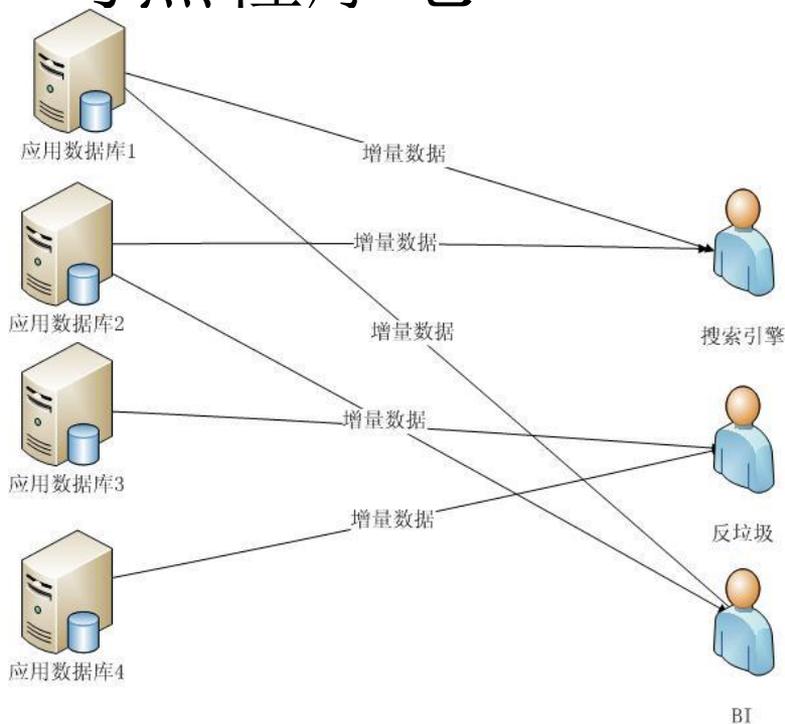
# 场景2

- 线上的服务器日志做下分析吧，可能有DDOS



# 场景3

- 要拿数据库增量数据，建索引、反攻击、做分析，写点程序吧

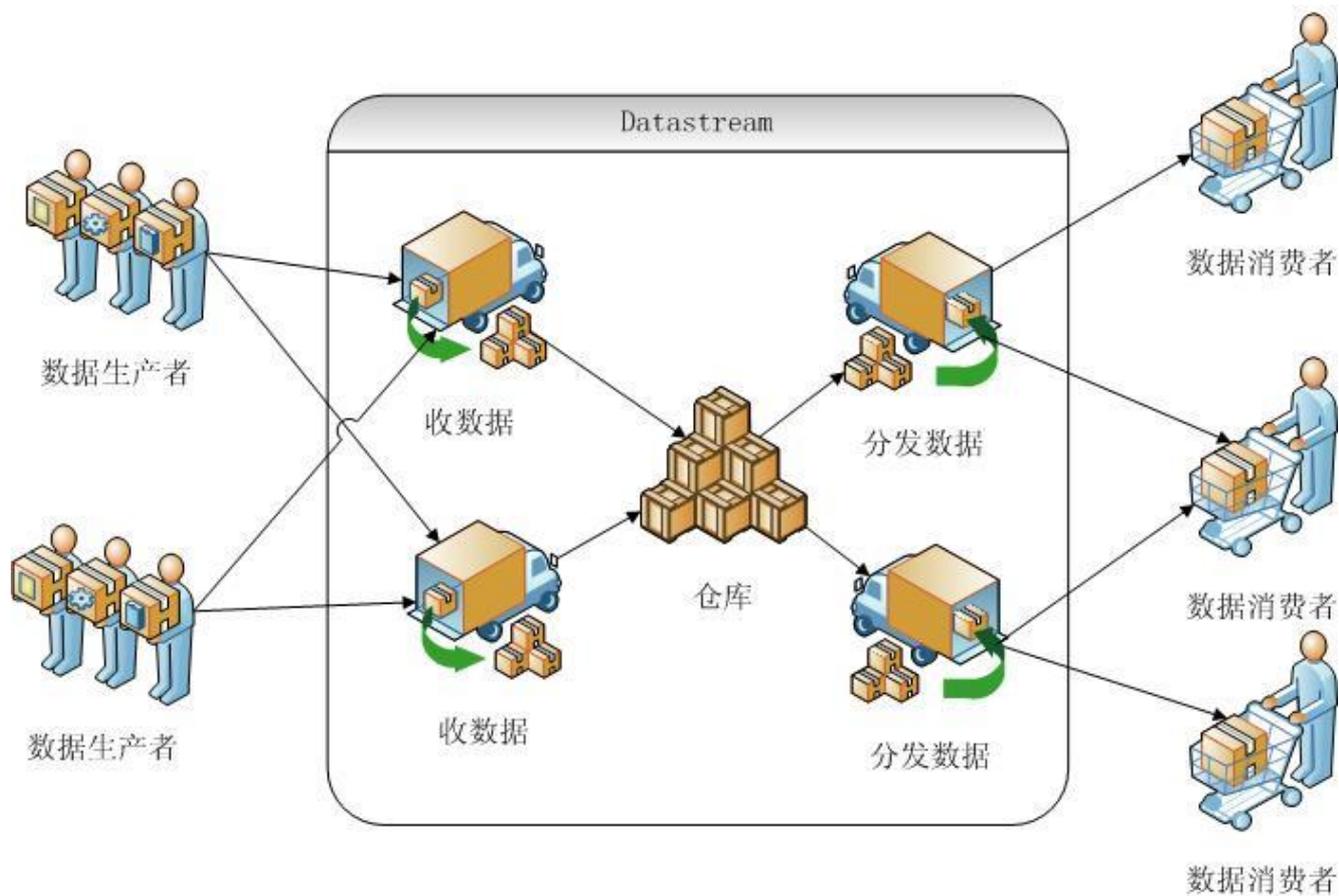


# 怎么办呢

- 问题:每个产品数据都不相同, 不同部门代码复用率低
- 后果: BI、反垃圾、搜索引擎的同学都要疯了
- 解决方案: 我们搞个统一的收集数据的平台吧

# Datastream应运而生

- 我们不生产数据，我们是数据的搬运工



# 大纲

- Datastream产生的原因
- Datastream的结构和特征
- Datastream关键技术点分享
- Datastream应用场景
- Datastream未来展望

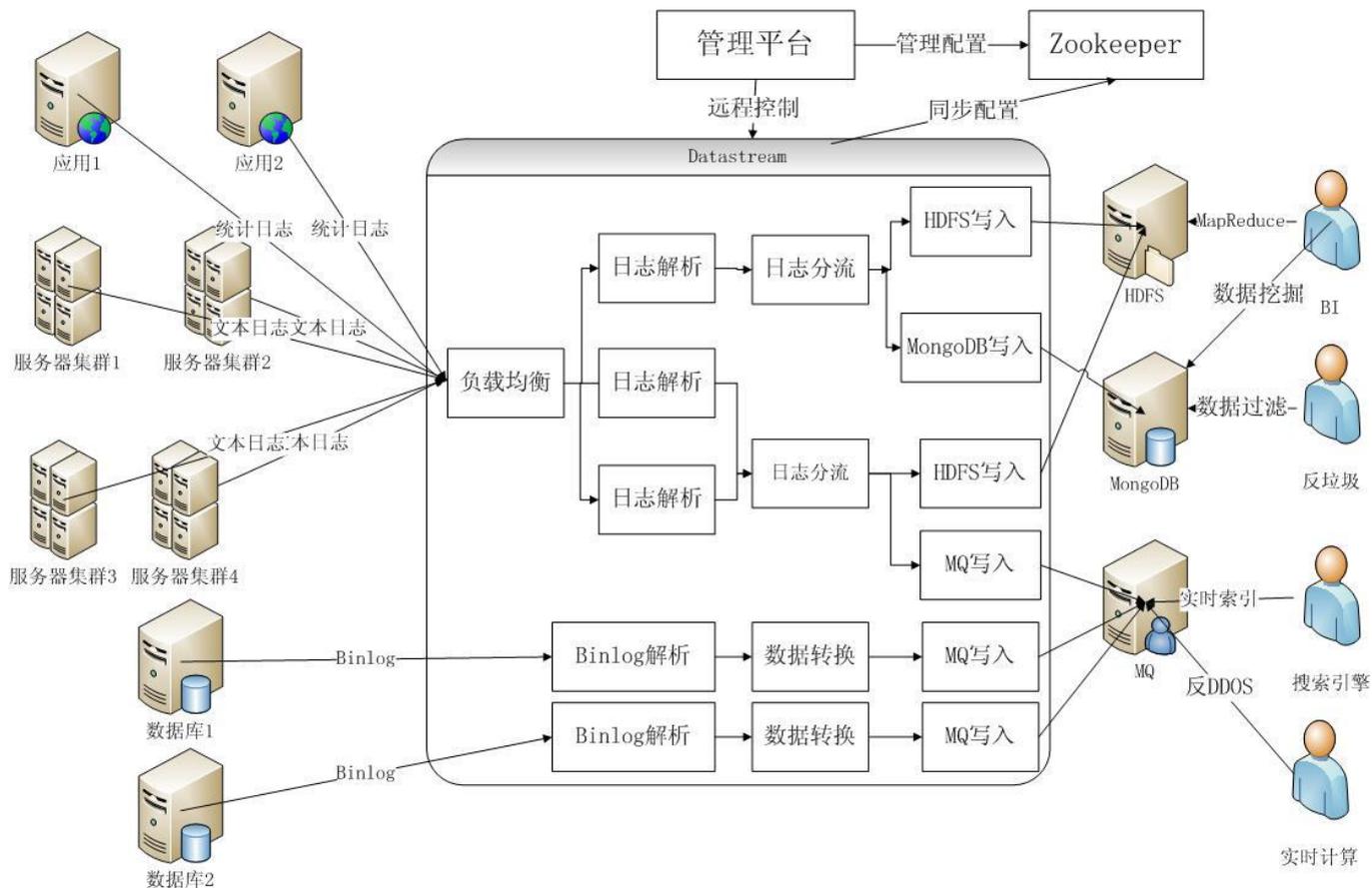
# 什么是Datastream

- Datastream是一个数据传输平台，用于连通数据源和终端用户的桥梁

他最大的作用就是：

- 让终端用户对数据获取方式透明
- 让终端用户对数据的格式透明

# 系统框架



# 系统特征

特性	说明
搬运节点线性扩展	集群中节点可动态增减，个别节点宕机不影响集群
数据来源多样化	可支持多种数据源，新数据源支持简单
数据存储终端多样化	可根据产品需要支持不同类型的数据存储终端
数据分流和复制	根据产品需要对数据进行分流和复制
输出数据格式统一	统一采用结构化JSON的数据输出格式，产品端提取数据简单
强大的管理平台	管理平台有各种强大的向导功能
远程控制	所有的Datastream客户端都通过管理平台统一做远程控制，无需登录服务器

# 大纲

- Datastream产生的原因
- Datastream的结构和特征
- **Datastream关键技术点分享**
- Datastream应用场景
- Datastream未来展望

# 关键技术点

- 一、异构数据源解析
- 二、数据分流
- 三、数据可靠传输
- 四、数据传输效率

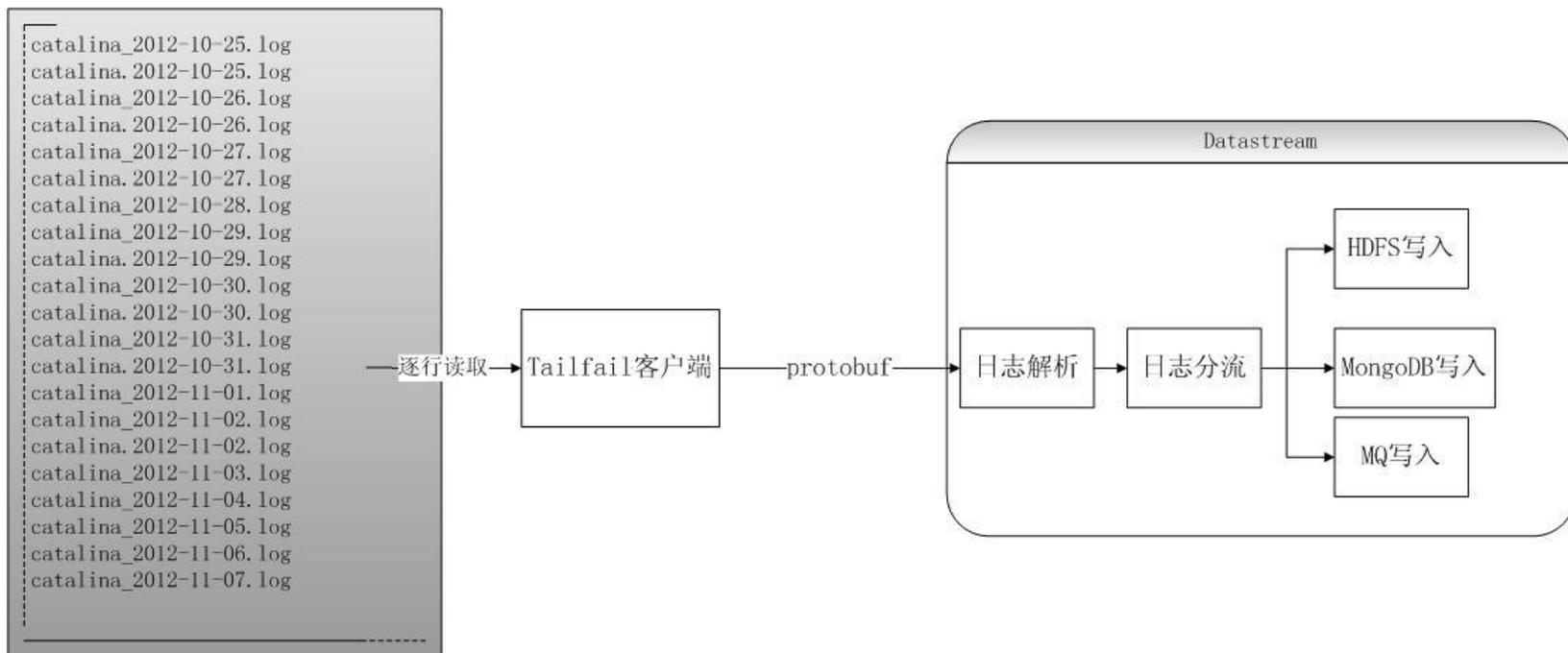
# 一、异构数据源解析

- Datastream从扩展性上来看能支持各种各样的数据源，目前支持一下数据源：
  - 文本文件
  - Syslog日志
  - Mysql binlog

# 文本日志来源

- 主要是应用服务器的日志文件

通过一个轻量级的agent来tail服务器本地日志并发送到datastream



# 文本日志解析的特点

- 支持各种滚动方式的日志
- 支持断点续传
- 轻量级，占用系统资源少
- 对应用无侵入
- 支持多行相关联日志输出为一行（例如java异常，结构化的JSON，xml等）

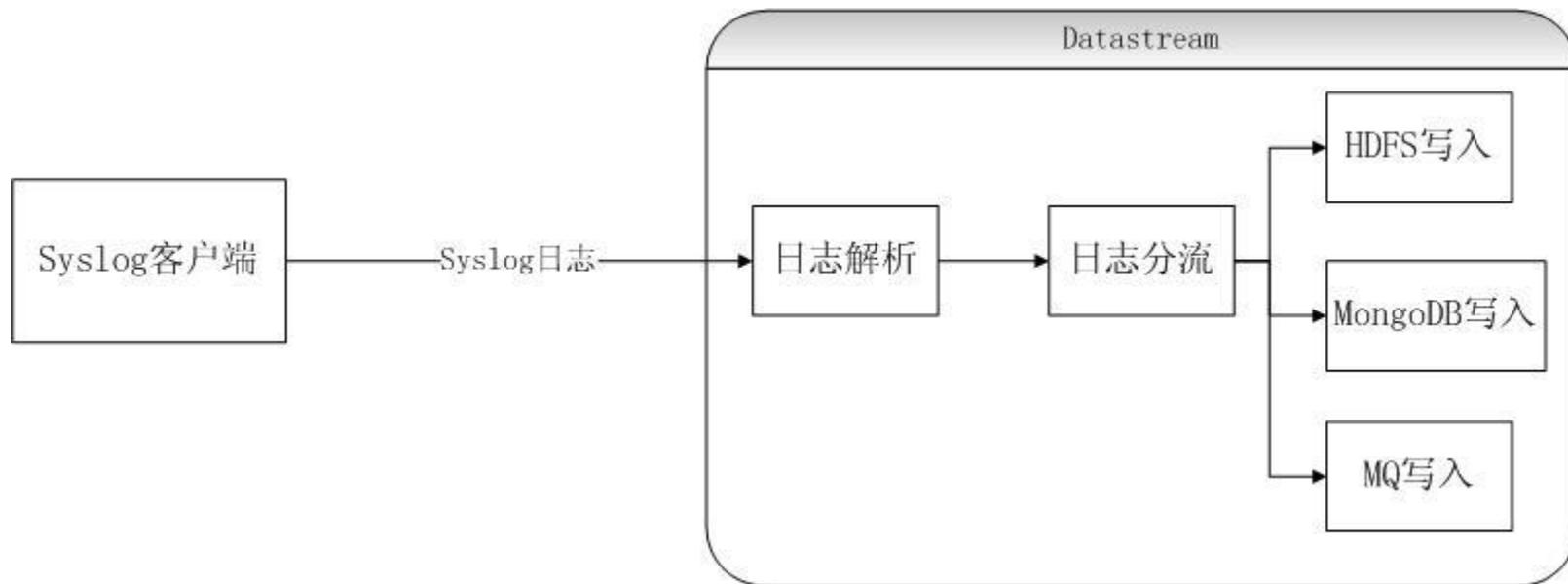
# tailFile多行解析的实例

```
2012-11-29 13:18:45,769 648222525 [catalina-exec-302] DEBUG - {
  "checkIfCompleted" : true,
  "checkIfResume" : true,
  "doUploadControl" : false,
  "file" : {
    "description" : "",
    "doc" : {
      "md5" : "e2c9eel8aal8fddbba538f567d7fb235",
      "size" : 10163640,
      "type" : 0
    },
    "fileName" : "qq2009preview_chs.exe",
    "incomplete" : {
      "local" : {
        "path" : ""
      },
      "ownerId" : 4000030838,
      "temporary" : null
    },
    "ownerId" : 4000030838,
    "parentId" : 485304124798,
    "type" : 0
  },
  "nameConflictPolicy" : 1,
  "renamePolicy" : 0
}
```

# Syslog 日志来源

- 这是一个通用的日志协议，不受编程语言限制，来源有
  - 服务器系统日志
  - 应用的重要日志
  - Apache,nginx等代理日志

# Syslog日志

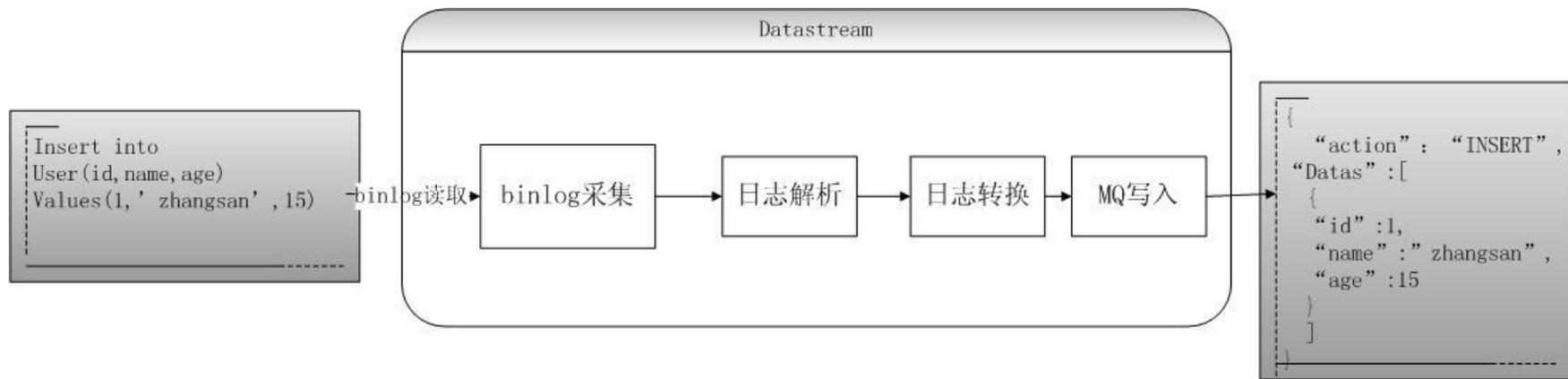


# Syslog 日志解析特征

- 平台无关
- 应用侵入性低

# Mysql binlog来源

- Mysql binlog是mysql增量数据的记录，对于实时计算有相当大的作用



# Mysql binlog解析关键技术

- 全自动化用户配置
- Binlog转换JSON
- 反向查询
- 事务支持
- 强顺序性保证

# 全自动化用户配置

- 全自动化加载数据库表结构，支持分布式数据库

<input checked="" type="checkbox"/>	字段名称	字段类型	注释	主键	均衡字段
<input checked="" type="checkbox"/>	id	BIGINT	笔记id	√	
<input checked="" type="checkbox"/>	user_id	BIGINT	用户id		
<input checked="" type="checkbox"/>	course_id	BIGINT	课程id		√
<input checked="" type="checkbox"/>	gmt_create	BIGINT	笔记创建时间		
<input checked="" type="checkbox"/>	gmt_modified	BIGINT	笔记修改时间		
<input checked="" type="checkbox"/>	title	VARCHAR	提问标题		
<input checked="" type="checkbox"/>	replied_num	SMALLINT	问答已回答数量		
<input checked="" type="checkbox"/>	best_reply_id	BIGINT	最佳答案id		
<input checked="" type="checkbox"/>	active_flag	SMALLINT	是否删除标记，0-删除，1-存在		
<input checked="" type="checkbox"/>	content	TEXT	问题补充		
<input checked="" type="checkbox"/>	browse_count	BIGINT	浏览数		
<input checked="" type="checkbox"/>	agree_count	BIGINT	问题顶的次数		
<input checked="" type="checkbox"/>	disagree_count	BIGINT	问题踩的次数		

# 二进制binlog转换为JSON

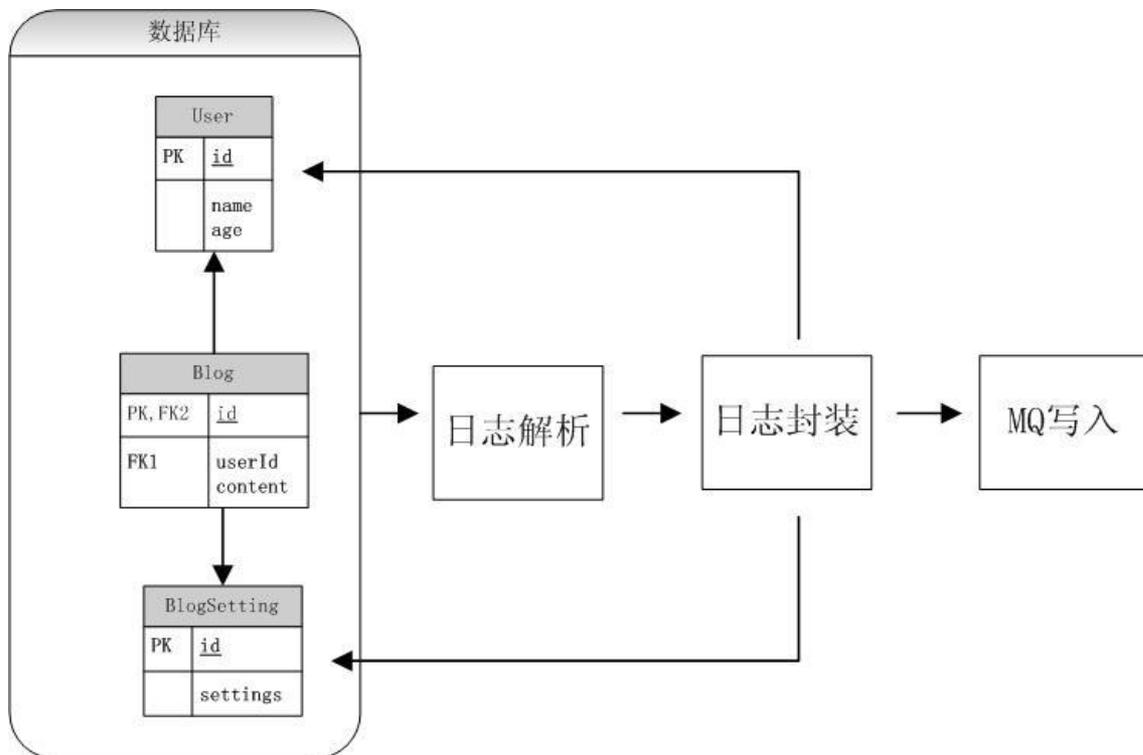
- 可以根据用户配置将数据库中的增量数据转换成结构化的JSON，供产品使用

属性名称	属性类型	数据库	表	表字段	是否监控	关键属性	是否发送	操作
id	NUMBER	edu	ask	id	是	是	是	X
user_id	NUMBER	edu	ask	user_id	是	是	是	X
course_id	NUMBER	edu	ask	course_id	是	是	是	X
gmt_create	NUMBER	edu	ask	gmt_create	是	是	是	X
title	STRING	edu	ask	title	是	是	是	X
replied_num	NUMBER	edu	ask	replied_num	是	是	是	X
best_reply_id	NUMBER	edu	ask	best_reply_id	是	是	是	X
active_flag	NUMBER	edu	ask	active_flag	是	是	是	X
content	STRING	edu	ask	content	是	是	是	X
labels	STRING	edu	ask	labels	是	是	是	X
title2	STRING	edu	ask	title	是	是	是	X
labels2	STRING	edu	ask	labels	是	是	是	X
browse_count	NUMBER	edu	ask	browse_count	是	是	是	X
agree_count	NUMBER	edu	ask	agree_count	是	是	是	X
disagree_count	NUMBER	edu	ask	disagree_count	是	是	是	X

# 反向查询

- 反向查询

有些应用（比如实时索引），需要变化的数据关联的其他数据构成完整数据



# 事务支持

- 支持mysql事务，一个事务内的操作会封装为一个在一个JSON内，同时保持事务的执行顺序
- 过大的事务会进行切割，避免内存消耗过大

# 强顺序性保证

- binlog的顺序如果出错，可能会导致后端产品最终数据状态的错误，因此datastream在处理Binlog时采取了强顺序性的保证
  - 每个mysql节点的事务操作顺序依次发送
  - 事务内操作保持顺序
  - MQ消费客户端的ack机制保证消费的顺序性和不遗漏消息
  - MQ消费客户端异常退出会重新获取上次未完成的消息

## 二、数据分流

- 日志分流规则的基础为数据标签
- 数据标签唯一标志一类数据的分流方式

名称

产品  \*

编码

过滤规则

添加分流规则

<b>HDFS</b> X	<b>MONGODB</b> X	<b>MQ</b> X
命名规则: cloud_rds_	数据库: datastream_1	服务器: db-48.photo.163.org
存储路径: /cloud/rds	存储集合: cloud.rds	队列名称: cloud.rds
归档时间: 1 小时	存储模式: safe	
归档大小: 不限制		

# 数据分流隔离机制

- 不同数据标签之前数据传输隔离，保证不同流量的产品之间传输速度不会相互影响
- 不同分流方式之间隔离，保证某个数据终端宕机或传输速度慢不会影响其他分流方式

## 三、数据可靠传输

- 持久化机制
  - 收到消息先做持久化后发送
- ACK机制
  - 确保后端模块已收到数据，否则重发
- 异常数据处理
  - 无法处理的异常数据保存在制定位置，可追溯来源

# 数据可靠传输

- 全面覆盖的监控程序
  - 监控所有进程的正确执行
  - 监控系统有无异常日志
  - 监控系统吞吐量是否不足有延迟
  - 监控产品数据流量异常波动
  - 监控产品有无不符合协议的日志

## 四、数据传输效率

- 分布式处理，可线性拓展
- 数据批量传输,批量ACK
- 采用thrift序列化机制
- 采用nio优化数据持久化效率
- 多线程处理日志并保证顺序性（如日志过滤）

# 大纲

- Datastream产生的原因
- Datastream的结构和特征
- Datastream关键技术点分享
- Datastream应用场景
- Datastream未来展望

# Datastream应用场景

- 一般产品需要数据做哪些操作呢？
  - 存储备份
  - 条件查询
  - 分析运算
  - 构建索引

Datastream都能满足

# Datastream应用场景

- Datastream目前可以支持以下应用场景：
  - 离线数据分析
  - 实时数据分析
  - 实时索引
  - 日志归档
  - 日志统计报表
  - 日志报警监控

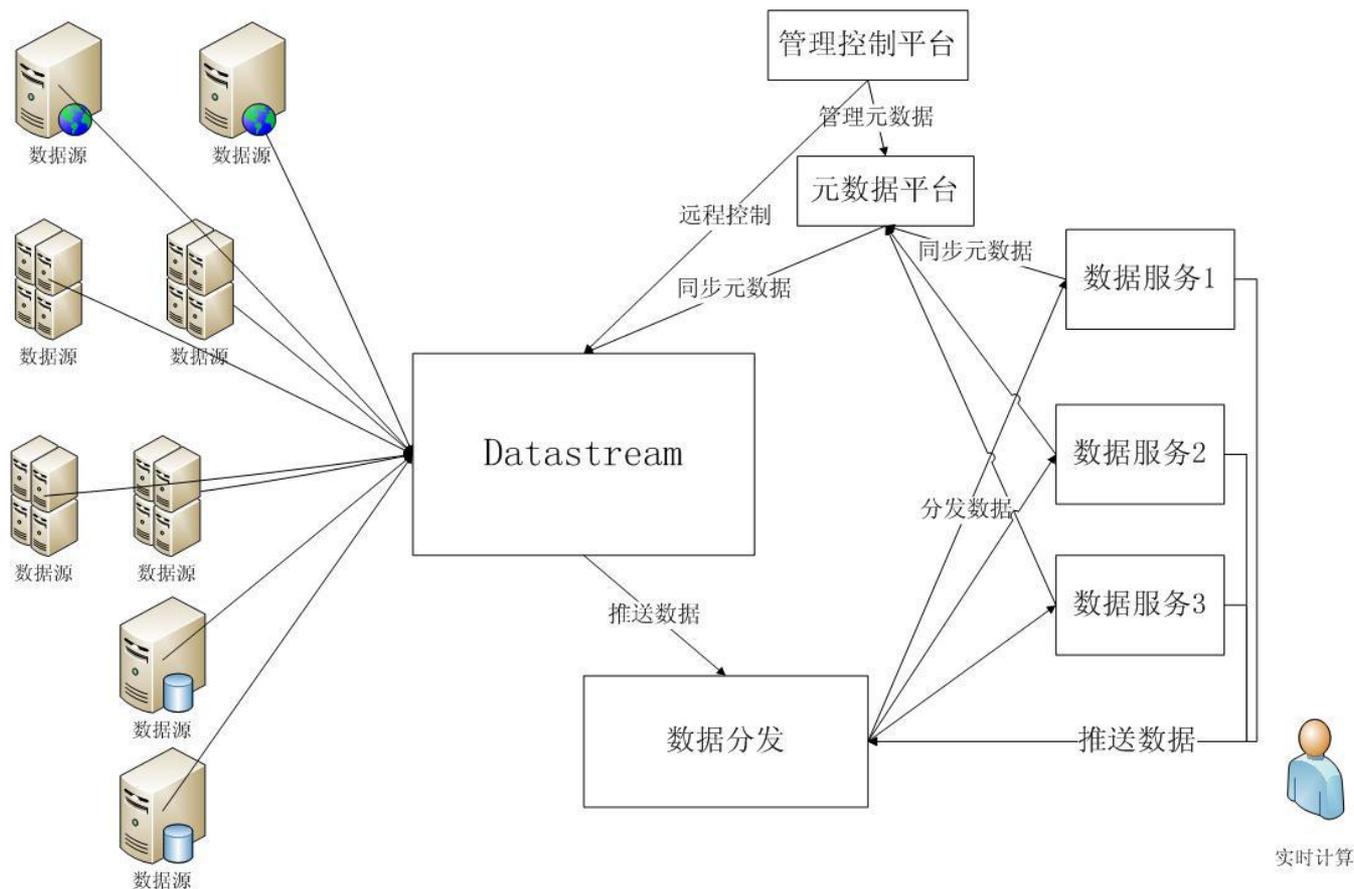
# 大纲

- Datastream产生的原因
- Datastream的结构和特征
- Datastream关键技术点分享
- Datastream应用场景
- Datastream未来展望

# Datastream未来展望

- Datastream项目将来会更好发挥搬运工角色，做一个效率更高，安全性更好的搬运工
- 后期会考虑实现如下的功能：
  - 数据追踪系统，提升数据安全性，类似快递
  - 更好地与后端数据服务整合
  - 非结构化日志转为结构化日志

# 未来的Datastream



# Q&A

